

3/PATS₁

09/914240

19.12.2000

JCU5 Rec'd PCT/PTO

24 AUG 2001

SNR scalable video encoding method and corresponding decoding method

FIELD OF THE INVENTION

The present invention relates to an encoding method for use in an SNR (signal-to-noise ratio) scalable video encoder, and to a corresponding decoding method.

5 BACKGROUND OF THE INVENTION

In video compression, scalability is the expected functionality to address the ever growing constraints of video transmission over heterogeneous networks (bandwidth, error rate...) in terms of varying receiver capabilities and demands (CPU, display size, application). It indeed allows a progressive transmission of information (in layers, or not) in order to provide a quality level of the reconstructed video sequence that is proportional to the amount of information that is taken of the bitstream.

Although they have not been initially designed to address these issues, current standards tried to upgrade their video coding schemes in order to include this functionality. In quality or SNR scalable compression schemes, temporal and spatial resolutions are kept the same, but the image quality is intended to vary depending on how much of the bitstream is decoded. In practice, most standards provide SNR scalability by means of a layered structure without giving up the classical single-scale scheme. The base layer (BL) is generally highly and efficiently compressed by a hybrid predictive encoding loop. The enhancement layer (EL) improves the quality of the compressed video signal by encoding the residual error (or prediction error), which is the difference between the original image and a reconstructed image. In MPEG-4 version 4, for instance, the EL uses DCT bit-planes to reencode this residual error.

The resulting scalability is however suboptimal for two main reasons. First, it is only based on an additional encoding of the prediction error and does not involve any refinement of the motion estimation and compensation processes. Moreover, it employs coding techniques like DCT that are not intrinsically designed to provide a progressive information transmission. For a more efficient scalable video coding, hierarchical strategies then appear to be promising candidates. The main idea is to design schemes that provide a generalized hierarchical representation of the information, opening the way to scalability.

Schematically, a simple hierarchical video coding scheme may be composed of several levels, each of which delivers a better-reconstructed image by means of a global refinement process (for instance, the hierarchy may use a pyramid composed of several image resolutions).

- 5 In parallel, hierarchical hybrid predictive coding schemes using non-block image representations like triangular meshes also constitute an interesting alternative. Meshes are well adapted to prediction error coding since they efficiently make the distinction between smooth regions and contours, well-compensated areas and occlusion regions. However, existing mesh-based methods generally encode the prediction error in a traditional way (block-based DCT, for example), i.e. by treating the error image as a whole picture
10 without using the mesh employed during the motion estimation and compensation stages. These methods suffer from a lack of flexibility, especially at low bit-rates, and do not provide an embedded bitstream.

- An alternative to DCT, better suited to low bit rate and tested in
15 MPEG-4, is based on the so-called Matching Pursuit (MP) algorithm, described for instance in "Matching pursuits with time-frequency dictionaries", by S.Mallat and Z. Zhang, IEEE Transactions on Signal Processing, vol.41, n°12, December 1993, pp.3397-3415. Indeed, MP is particularly well suited to the progressive texture encoding of arbitrarily shaped objects. Moreover, an intrinsic way of providing SNR scalability with MP is through the number of
20 encoded "atoms". MP naturally achieves scalability by encoding the motion prediction error in decreasing order of energy. The procedure is iteratively applied until either the bit budget is exhausted or the distortion falls down below a prespecified threshold. The granularity of MP is the coding cost of one atom, that is approximately 20 bits.

- In the European patent application filed on December 28th, 1999, under the
25 filing number 99403307.4 (PHF99627), an MP prediction error coding method has been included inside a hierarchical mesh-based video coding scheme, which allows to benefit from the triangular mesh advantages concerning spatial adaptability, deformation capacity, compact and robust motion estimation even at low bit rates. The mesh hierarchy of this scheme is obtained through a coarse-to-fine strategy, beginning at the first level with a coarse
30 regular triangular mesh, and which then includes a mesh refinement process locally subdividing triangles of the current level where the prediction error signal is still important after motion compensation (the new mesh is taken as the input for the following level). Based on the MP algorithm, this method benefits from the mesh characteristics, while being especially designed to match the triangle support. Given any selected triangle, the issue is

3/3

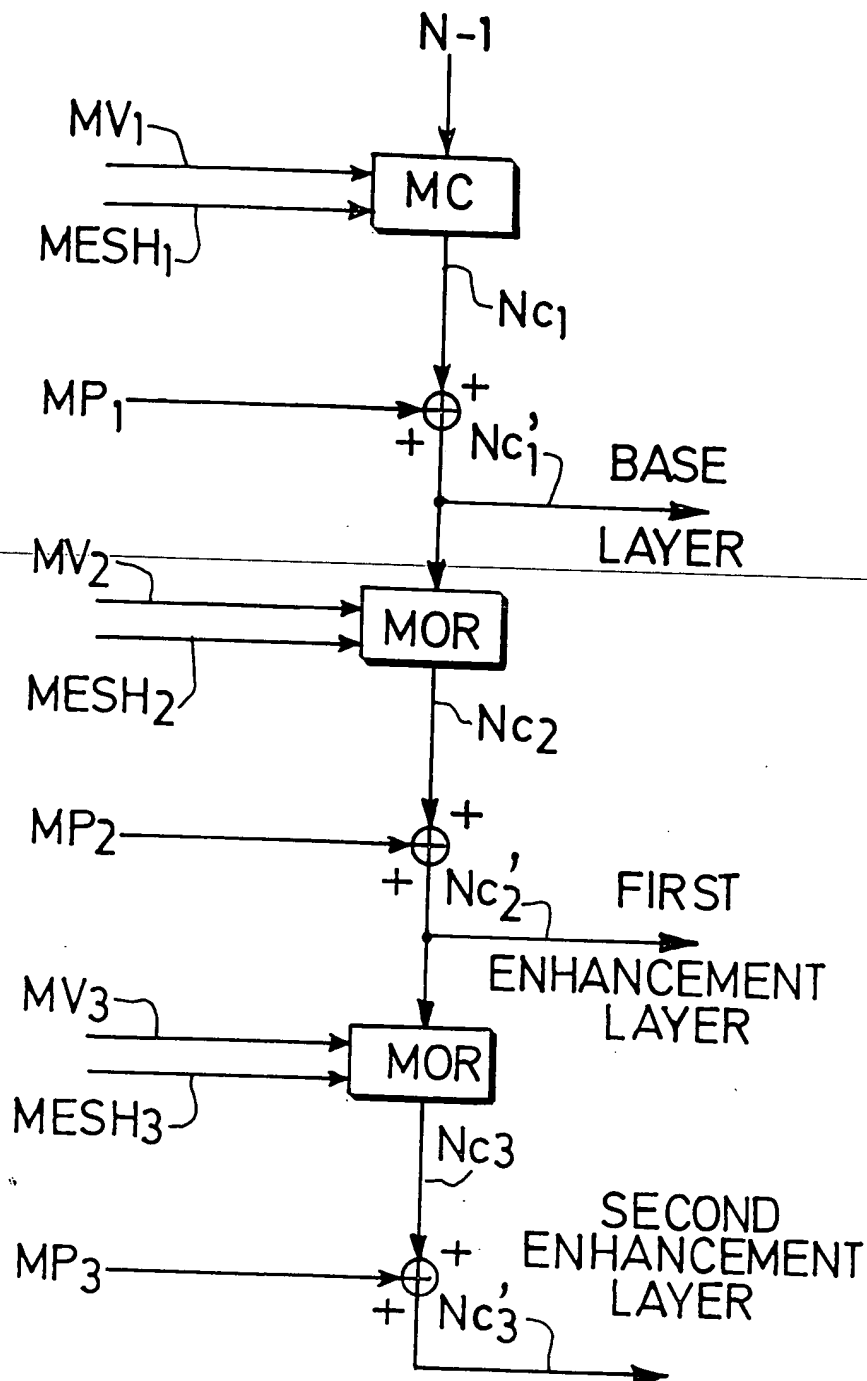


FIG.3